CSCanada
Canadian Social Science
Vol. 13, No. 7, 2017, pp. 1-5
DOI:10.3968/9698

ISSN 1712-8056[Print]
ISSN 1923-6697[Online]
www.cscanada.net
www.cscanada.org

Application of Post-Editing in Foreign Language Teaching: Problems and Challenges

SUN Dongyun[a],*

[a]Lecturer, College of Foreign Languages and Literatures, Fudan University, Shanghai, China.
*Corresponding author.

Supported by Ministry of Education of China (15YJC740071).

Received 10 April 2017; accepted 19 June 2017
Published online 26 July 2017

Abstract

Introducing the basic concepts of machine translation and post-editing (PE), this paper points out that PE is growing rapidly in the language service industry, and then surveys the status quo of its application in translator training, foreign language learning, development of translation techniques, and writing. Based on this survey, this paper shows various problems of PE in language learning in terms of learner orientation, ethics and translators' status. It is further pointed out that though researches in PE are immature and inadequate, its potential impact on learners deserves greater attention. Instead of simply boycotting PE, foreign language teachers and translation trainers should learn more about it in order to take countermeasures proactively.

Key words: Machine translation; Post-editing; Professional translation teaching; Foreign language teaching

Sun, D. Y. (2017). Application of Post-Editing in Foreign Language Teaching: Problems and Challenges. *Canadian Social Science, 13*(7), 1-5. Available from: http://www.cscanada.net/index.php/css/article/view/9698
DOI: http://dx.doi.org/10.3968/9698

INTRODUCTION

A. Definition of Post-Editing

Post-editing (PE) is a concept that is closely related to machine translation (MT). It "involves correcting the translation output generated by the machine translation system, a task performed by the human editor or translator in order to bring the text to a certain pre-determined standard in terms of language style and appropriate use" (Quah, 2006, p.11). Since fully-automatic high-quality machine translation (FAHQMT） is not feasible at present, post-editing MT outputs to a usable degree is a viable solution to the commercialization of MT systems.

A distinction should be made between human-aided/assisted machine translation (HAMT) machine-aided/assisted human translation (MAHT) (Bowker, 2002, p.4) with respect to discussions on translation technology. In the former model, MT plays a major role while human translators perform pre-editing or post-editing to make results of usable. In the latter case, human translators play a major role while computers facilitate the process with translation memory, terminology management, word-processing, dictionary lookup and referencing, etc., with MT excluded. The latter is traditionally regarded as the domain of computer-aided translation (CAT) (e.g. Bowker, 2002) but is now often expanded to include MT post-editing.

B. Significance of PE

Vasconcellos and Bostad (1992, p.68) believes that PE as a specialized skill should seek to retain认the MT outputs and does not make modifications unless necessary. Therefore post-editing MT outputs is technically a feasible way of saving labor costs and improving efficiency. Its significance can be described in the following aspects:

First, if PE is more labor-efficient than human translation, it is a cost efficient solution to commercial translation. Koehn (2009, p.23) believes that PE is cost-efficient if "the efforts of post-editing is less than the effort of translation from scratch" and that "[i]f the machine translation system reliably brings the meaning across, the post-editor does not need to know the foreign input language." Though such an idealized scenario is unattainable at present, one can reasonably imagine the

expediency of post-editing MT outputs by monolingual speakers in lieu of the minority language professionals that are hard to find. In fact, with the development of data-driven MT such as statistical machine translation and neural machine translation complemented by domain-specific data, the quality of MT outputs has improved significantly over the past decade. Garcia (2010, 2011) conducted researches of post-editing uncustomized Google MT engine outputs by professional translators and translation learners, and results showed that PE yields comparable or even superior performance in both time and quality to human translations in both cases.

Second, PE is often used in benchmarking the quality of MT outputs. Despite the prevalence of automatic scoring algorithms represented by BLEU and NIST, many scholars realized the inherent defects of these algorithms as they are often a far cry from human intuition. In view of this, the Human Translation Error Rate (HTER) scheme was incorporated in the development of DARPA MT system. Here HTER is defined as "the number of editing steps divided by the number of words in the acceptable translation," for "a human annotator has to find the minimum number of insertions, deletions, substitutions, and shifts to convert the system output into an acceptable translation." (Koehn, 2009, p.238) Of course, the number of steps alone cannot fully reflect the actual level of difficulty of PE, and thus the time spent on human PE should also be taken into consideration.

C. Recent Developments of PE

Despite the controversy over the usability of MT outputs, with the introduction of Google Neural Machine Translation (GNMT) in late 2016, the quality and readability of MT have improved significantly. In some domains, it only takes minimal efforts of PE to make the translation publishable. Therefore MT is no longer reserved for gisting only, but is on the verge of commercialization in the language service industry. The other advantage of statistical machine translation (SMT) and neural machine translation (NMT) is that they require no linguistic knowledge, reduces the human resources costs, are trained with human translations independent from the language pair, and easy to build and maintain. The unique advantage of NMT is that it can handle long-distance reordering and word agreement to achieve good fluency with a relatively small dataset (Pouliquen, 2016). Just as Langlais and Carl (2004) point out, STM coupled with domain-specific data can yield superior quality outputs. Pouliquen (2016) reported superior quality of World Intellectual Property Organization (WIPO) MT to Google MT in both cases of SMT and NMT, as evidenced by BLEU scores trained with patent data 2016 (see Figure 1).

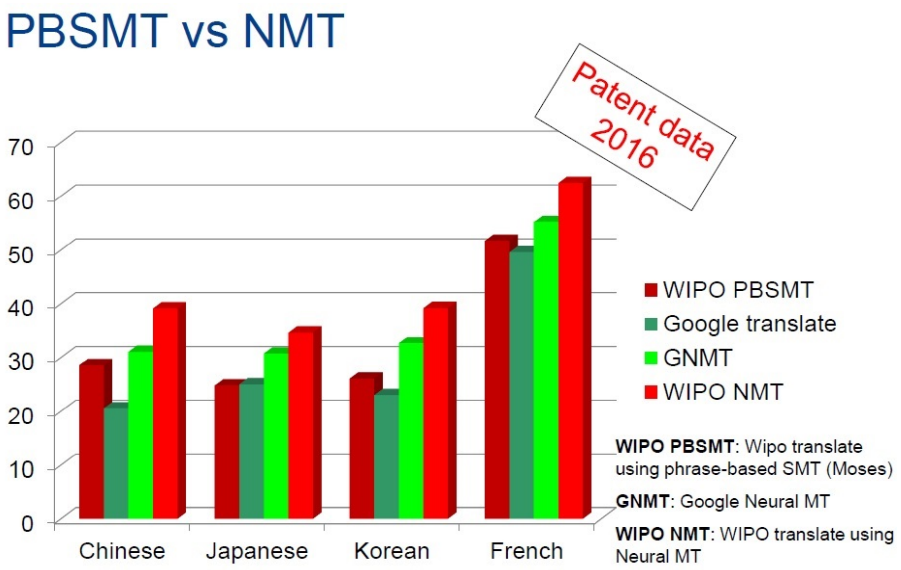

Figure 1
BLEU Scores of WIPO MT Systems vs. Google SMT and NMT (Pouliquen, 2016)

With the growing popularity of MT, an increasing number of players joined the game, including Google, Amazon, Facebook, WIPO and some of the leading IT companies in China such as Alibaba, Tencent, Sogou, Youdao and iFlytek. Post-editing outputs from these MT engines will grow exponentially and gain a bigger share in the language service industry in the foreseeable future.

1. THE APPLICATION OF PE IN FOREIGN LANGUAGE TEACHING

PE has been widely used in translation practice, and has attracted the attention of translation teachers and language teachers alike, who try to use post-editing to improve the efficiency of translation and language teaching. At present, PE is mainly used in the following areas:

1.1 Translator Training and Program Development

The application of MT in translator training has been underway for decades. The undergraduate program Machine Translation was established in Hebei Normal University in 1995, and their texbook *A Concise Course Book of Machine Translation* (2009) was published by Shanghai Foreign Language Education Press. In Shih's manuscript *Computer-Aided Translation MT & TM* (2004), nearly half of the volume was devoted to the discussion of PE skills and teaching methods. Somers (2003) introduces the basic principles of machine translation, the existing commercial machine translation systems, and skills of post-editing.

In recent years, with the development of translator training programs in our country, the Master of Translation and Interpreting (MTI) programs have been sprouting all over China. Courses on computer-aided translation (CAT) and translation technology have thus received unprecedented attention. Many MTI programs have now been equipped with CAT labs. While Bowker (2002) and Qian (2011) did not incorporate MT into the CAT curriculum, Wang (2012, p.58) introduced MT as a course module for the CAT master program of Peking University, however, no post-editing skills were mentioned in the program description. With MT making its way into the CAT or localization programs in University of International Business and Economics and Beijing Language and Culture University, it has become an integral part in translator training.

1.2 PE as a Tool for Computer-Assisted Language Learning (CALL)

Niño (2009, pp.241-242) pointed out that MT can be used as a bad example for foreign language learning, and its outputs can be used by language learners for learning error correction. Somers (2003, p.327) also stated that post-editing MT errors can help learners "learn about subtle language differences" and "strengthen learners' understanding of mother tongue and second language grammar and style." Niño (2004) provides students with an error correction task of post-editing an unedited MT translation into Spanish from an English text of general topic. Experiments show that this activity improves learners' linguistic and translation skills in both languages.

1.3 PE as a Tool for Learning Translation Skills

Shei and Pein (2002, p.323) designed a novel translation method tutoring system with MT and PE. First, they divided translation methods into a) verbatim translation, b) literal translation, c) semantic translation and d) communicative translation. Based on the cascade of deviations from the form of the source text, they devised a Translation Method Tutor (TMT) system, with its corpus composed of four different translations, each of which corresponding to the following translation method:

a) MT output (corresponding to "verbatim translation");
b) MT with post-editing (corresponding to "literal translation");
c) Text imported from bitext alignment (corresponding to "semantic translation"), and
d) Manual revision of the text of semantic translation, rending it to conform to communicative translation.

The basic process goes like this: students translate the source text sentence by sentence. When they upload the texts into the TMT system, it automatically compares this translation with existing translations in the system to yield the closest match. The match results and its corresponding translation method are automatically provided to students as feedback. This system rides on the literal faithfulness of MT but uses PE results as examples of literal translation. However, with the development of NMT, the impression of "verbatim translation" of MT may need to be updated.

1.4 PE as English Writing Assistant

Regular English writing courses require students to write in English from their mind instead of translating from their mother tongue. However, Yang and Xiao (2012) stated that in some Tier-2 universities in China where students' English competence is limited, it is very hard for students to write their bachelor thesis in English, and consequently they had to resort to PE of MT outputs of their Chinese manuscripts. Garcia and Pena (2011) prove that beginners of foreign languages can effectively improve the efficiency and quality of writing in foreign languages. By contrast, screen videos captured during the writing process show that writing directly in the foreign language requires more cognitive efforts, with more pauses and concentration on the part of subjects.

2. LIMITATIONS OF PE IN FOREIGN LANGUAGE TEACHING

2.1 Lack of PE-Oriented Translation Teaching Model

Despite the growing popularity of PE in the translation industry, as evidenced by the integration of CAT tools with MT, there is a lack of consensus on how to teach PE in translator training. Although Shih (2004) discussed error analysis of MT outputs and PE strategies while Luo and Li (2012) discussed the error typology of Huajian MT outputs as well as PE techniques, the TransWhiz MT outputs are so drastically different from those of Huajian MT engine that their PE techniques are hardly compatible. It also has to note that both TransWhiz and Huajian are rule-based MT systems, but Google NMT has far surpassed these two MT engines in performance and accuracy. This reveals one of the dilemmas of PE research: since MT technologies are developing by leaps and bounds, by the time the PE guidelines for one MT

system are developed, the market might have already opted for another new MT system.

2.2 PE Changed the Subjectivity of Translators

For decades, the translation academia is hostile to the development of MT, thinking that human translators can never be replaced by MT, and ridiculing the poor quality of MT outputs. However, with the significant progress of MT in the recent decade or two, MT has become a viable option in the language service industry, and MT post-editor is emerging as a new profession.

However, who is indeed the dominant player in the activity of translation? With respect to traditional CAT, all decision-making processes are up to the discretion of the translator, with the computer playing a subordinate role. Hence the incessant controversy over the scope of CAT, ranging from both machine-aided human translation (MAHT) and human-aided machine translation (HAMT) by Hutchins and Somers (1992, p.148) to Bowker (2002) and Quah (2006)'s exclusion of MT and HAMT. With translators relegated to post-editors serving the computer, who is the real author of the target text? If the computer can be compared to an author submitting his manuscript to the editorial office while the translator is compared to the editor, then who does the intellectual property right go to? If translators' subjectivity is further curbed rather than enhanced by technology, how can they be motivated to contribute their ingenuity to translation quality assurance?

2.3 The Orientation of PE in Foreign Language Learning

The use of PE in foreign language writing is a more controversial topic. Although it is undeniable that MT plus PE can enable underachievers to write in a foreign language quickly with satisfactory quality, this "shortcut" will possibly undermine the writing habit and competence. Will the abuse of PE lead to overdue reliance on technology rather than on language learning? This issue deserves ongoing concern from teachers and researchers.

From the perspective of translation teaching, PE also poses an ethical dilemma. If learners do their homework with MT plus PE, does this conduct amount to plagiarism? Bendana and Melby (2012, p.45) believes that SMT is "essentially a tool for massive sophisticated plagiarism." Though this quip conveys the metaphor of the underlying statistical mechanism of leveraging fragmentary language data into reordered target texts, who do we attribute the post-edited MT output? To the learner or to the translator? Or to the author of each language fragment?

The use of PE could also undermine the development of translation competence. The teaching of translation skills, especially the skills of sentence recasting, has been an integral part of translation teaching. However, the training of PE skills could change the bias of translation teaching. Of PE skills and sentence translation skills, which is more important for beginners? If the latter counts more, then how to balance the relationship between the two?

CONCLUSION

As mentioned above, teachers and researchers in foreign language teaching have yet to learn about MT and PE, still less to use them. With regard to translator training, institutions rarely teach post-editing skills. Literature hitherto available shows that the teaching mode of PE is not yet mature, with different MT systems yielding different translation, thereby requiring different PE guidelines. Therefore from the current point of view, PE has not yet exerted any substantive impact on foreign language learning. But with the rise of post-editing in the language service industry, its repercussions are bound to spread to the education sector. The majority of translation teachers and foreign language teachers should be sensitive and concerned about this, and actively embrace new technology in order to adapt to the new changes of the era.

For the time being, the general view of the foreign language community is that the quality of MT is not good enough to cause any real impact on translation and language teaching. However, if the quality of MT continues to improve, then the learners will probably ignore the teacher's repeated admonition by secretly using PE for finishing the foreign language writing tasks. In the course of the development of foreign language teaching, similar experiences are not uncommon: 20 years ago, with the popularity of personal computers, more and more users use typing instead of handwriting. Many language teachers worried that the learners would be much less competent in handwriting. More than a decade ago, electronic dictionaries began to catch on. Regardless of the teacher's repeated warning, students began to use electronic dictionaries in lieu of the print dictionaries. Although some of the concerns have now become reality, for example, with the MS-Word spelling check application, many learners and even native speakers do ignore the good habits of spelling, which thus made them less and less competent in spelling. The trend is difficult to reverse. How to plan ahead by enabling learners with advanced tools and technology? Foreign language and translation teachers should play a major task.

At present. MT plus PE could expectedly cause some real impact on the translation industry. Moreover, its influence on translation teaching should not be overlooked. The author predicts that the continuous improvement of MT technology contributes to automation in some specific areas and specific text types such as software interface and product brochures. Moreover, translator training will also develop towards hierarchical differentiation, with less proficient language learners

doing the humdrum PE task of translation on the one end of the spectrum while highly competent translators dominating the higher-end market with greater added value. It is worth pondering the question of how to deal with this challenge amid the "gold rush" towards the professional education of translators and interpreters.

REFERENCES

Bendana, L., & Melby, A. K. (2012). *Almost everything you ever wanted to know about translation.* Retrieved November 20 from http://multi-languages.com/materials/everything_you_ever_wanted_to_know_about_translation_melby_bendana.pdf

Bowker, L. (2002). *Computer-aided translation technology: A practical introduction*. Ottawa: University of Ottawa Press.

Garcia, I. (2010). Is machine translation ready yet? *Target. International Journal of Translation Studies*, *22*(1), 7-21.

Garcia, I. (2011). Translating by post-editing: Is it the way forward? *Machine Translation*, *25*(3), 217-237.

Garcia, I., & Pena, M. (2011). Machine translation-assisted language learning: Writing for beginners. *Computer Assisted Language Learning, 17,* 471-487.

Hutchins, W. J., & Somers, H. L. (1992). *An introduction to machine translation* (Vol. 362). London: Academic Press.

Koehn, P. (2009). *Statistical machine translation*. Cambridge: Cambridge University Press.

Langlais, P., & Carl, M. (2004). General-purpose statistical translation engine and domain specific texts: Would it work? *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, *10*(1), 131-153.

Li, Z., & Meng, J. (2009). *A concise course book of machine translation*. Shanghai: Shanghai Foreign Language Education Press.

Luo, J., & Li, M. (2012). Error analysis of machine translation output. *Chinese Translators Journal, 5,* 84-89.

Niño, A. (2004). *Recycling MT: A course on FL writing via MT post-editing* (pp.179-187). Paper presented at CLUK (Computational Linguistics United Kingdom 7th Annual Research Colloquium), 6th and 7th January 2004 in the University of Birmingham, UK.

Niño, A. (2009). Machine translation in foreign language learning: Language learners' and tutors' perceptions of its advantages and disadvantages. *ReCALL*, *21*(2), 241-258.

Pouliquen, B. (2016). *Practical use of machine translation in international organizations*. Presentation at ICON 2016, Dec. 2016, Varanasi. Retrieved 2017, June 1 from http://ltrc.iiit.ac.in/icon2016/proceedings/icon2016/pdf/W16-6301.pdf

Qian, D. (2011). *Computer-aided translation*. Beijing: Foreign Language Teaching and Research Press.

Quah, C. K. (2006). *Translation and technology.* Hampshire/ New York: Palgrave Macmillan.

Shei, C., & Pein, H. (2002). Computer-assisted teaching of translation methods. *Literary and Linguistic Computing*, (3), 323-344.

Shih, C. (2004). Computer-aided translation MT&TM. Taipei: Bookman Publishing Company.

Somers, H. L. (2003). MT in the classroom. In H. Somers (Ed.), *Computers and translation: A translator's guide* (pp.319-340). Amsterdam/Philadelphia: John Benjamins.

Vasconcellos, M., & Bostad, D. (1992). Machine translation in a high-volume translation environment. In J. Newton (Ed.). *Computers in translation: A practical appraisal* (pp.58-77). London and New York: Routledge.

Wang, H. (2012). Teaching practice of translation technology in the context of information age. *Chinese Translators Journal, 3,* 57-62.

Yang, J., & Xiao S. (2012). *Revising Chinese-English machine translation with online corpora.* Paper presented at 2012 International Conference on Application of Technology in Translation and Foreign Language Teaching. Shanghai: Shanghai University of Technology.